A family of rules for parameter choice in Tikhonov regularization of ill-posed problems with inexact noise level

U. Hämarik, R. Palm, T. Raus

University of Tartu, Estonia

International Conference on Scientific Computing

S. Margherita di Pula, Sardinia, Italy

October 14, 2011

Contents

- We consider linear ill-posed problems

$$Ax = y_*, \qquad y_* \in \mathcal{R}(A),$$

where $A\colon X \to Y$ is a linear continuous operator between Hilbert spaces. The range $\mathcal{R}(A)$ may be non-closed and the kernel $\mathcal{N}(A)$ may be non-trivial.

- Assume that instead of exact data y_* only its approximation y is available.
- For approximation of the minimum norm solution x_* of the problem $Ax = y_*$ we use the Tikhonov regularization method

$$x_\alpha = (\alpha I + A^* A)^{-1} A^* y.$$

Information about noise level

- In the following we consider three cases of knowledge about noise level for $\|y - y_*\|$:
  - Case 1: exact noise level δ: $\|y - y_*\| \leq \delta$.
  - Case 2: no information about $\|y - y_*\|$.
  - Case 3: approximate noise level: given is δ but it is not known whether the inequality $\|y - y_*\| \leq \delta$ holds or not. For example, it may be known that with high probability $\delta/\|y - y_*\| \in [1/10, 10]$. This very useful information should be used for choice of $\alpha = \alpha(\delta)$.
- Choice of regularization parameter α.
  - Rules for the Case 1 (discrepancy principle, etc.) need exact noise level: rules fail for very small underestimation of the noise level and give large error $\|x_\alpha - x_*\|$ already for 10% overestimation.
  - Rules for the Case 2 do not guarantee the convergence $x_\alpha \to x_*$ for $\|y - y_*\| \to 0$.
  - Our rules for the Case 3 guarantee $x_\alpha \to x_*$ as $\delta \to 0$, if $\lim_{\delta \to 0} \frac{\|y - y_*\|}{\delta} \leq \text{const}$.

Parameter choice rules for the case of exact noise level

- Discrepancy principle (D): α_{D} is the solution of $d_{\mathrm{D}}(\alpha) := \|Ax_\alpha - y\| = C\delta,\ C \geq 1$.
- Monotone error rule (ME):

$$d_{\mathrm{ME}}(\alpha) := \frac{\|B_\alpha(Ax_\alpha - y)\|^2}{\|B_\alpha^2(Ax_\alpha - y)\|} = \delta,$$

$$B_\alpha = \sqrt{\alpha}(\alpha I + AA^*)^{-1/2}.$$

Family of rules for parameter choice

Fix q, l, k such that $3/2 \le q < \infty$, $l \ge 0$, $k \ge l/q$; $2q$, $2k$, $2l \in \mathbb{N}$.
Choose $\alpha = \alpha(\delta)$ as the largest solution of

$$d(\alpha \mid q, l, k) := \frac{\kappa(\alpha) \| D_\alpha^k B_\alpha (Ax_\alpha - y)\|^{q/(q-1)}}{\| D_\alpha^l B_\alpha^{2q-2}(Ax_\alpha - y)\|^{1/(q-1)}} = b\delta,$$

where $B_\alpha = \sqrt{\alpha}(\alpha I + AA^*)^{-1/2}$, $D_\alpha = \alpha^{-1} AA^* B_\alpha^2$,

$$\kappa(\alpha) = \begin{cases} 1, & \text{if } k = l/q, \\ (1 + \alpha \|A\|^{-2})^{\frac{kq - l + q/2}{q-1}}, & \text{if } k > l/q, \end{cases} \tag{1}$$

$$\downarrow \ \alpha \to 0 \qquad\qquad 1 \tag{2}$$

$$b \approx \left(\frac{3}{2}\right)^{\frac{3}{2}} \frac{k^k}{(k+3/2)^{k+3/2}} \left(\frac{k^k (l+3/2)^{l+3/2}}{l^l (k+3/2)^{k+3/2}}\right)^{\frac{1}{q-1}}. \tag{3}$$

Denote this rule by R(q, l, k).

Examples of this family of rules

- Modified discrepancy principle (Raus 1985, Gfrerer 1987): $q = 3/2$, $l = k = 0$
- Monotone error rule (Tautenhahn 1998): $q = 2$, $l = k = 0$
- Rule R1 (Raus 1992): $q = 3/2$, $k = l > 0$
- Balancing principle (Mathé, Pereverzev 2003) can be considered as an approximate variant of rule R1 with $k = 1/2$.

Existence of solution for family of rules

1. If $k > l/q$, then the equation $d(\alpha \mid q, l, k) = b\delta$ has a solution for every $b = \text{const} > 0$, because $\lim_{\alpha\to\infty} d(\alpha \mid q, l, k) = \infty$ and $\lim_{\alpha\to 0} d(\alpha \mid q, l, k) = 0$.
2. If $k = l/q$, then the solution of the equation $d(\alpha \mid q, l, k) = b\delta$ exists, if $b \geq b_0(q, l, k)$ and $\|y - y_*\| \leq \delta$.

Convergence and stability

- **Convergence.** Let $k \geq l/q$. Let the parameter $\alpha = \alpha(\delta)$ be the solution of the equation $d(\alpha \mid q, l, k) = b\delta$, $b > b_0(q, l, k)$. If $\|y - y_*\| \leq \delta$, then $\|x_\alpha - x_*\| \to 0$ $(\delta \to 0)$.
- **Stability** (with respect to the inaccuracy of the noise level). Let $k > l/q$. Let the parameter $\alpha(\delta)$ be the **largest** solution of the equation $d(\alpha \mid q, l, k) = b\delta$. If $\frac{\|y-y_*\|}{\delta} \leq c = \text{const}$ in the process $\delta \to 0$, then $\|x_\alpha - x_*\| \to 0$ $(\delta \to 0)$.

Quasioptimality

Let $l/q \leq k \leq l \leq q/2$. Let the parameter $\alpha(\delta)$ be the **smallest** solution of the equation $d(\alpha \mid q, l, k) = b\delta$. Then the rule is ***quasioptimal***:

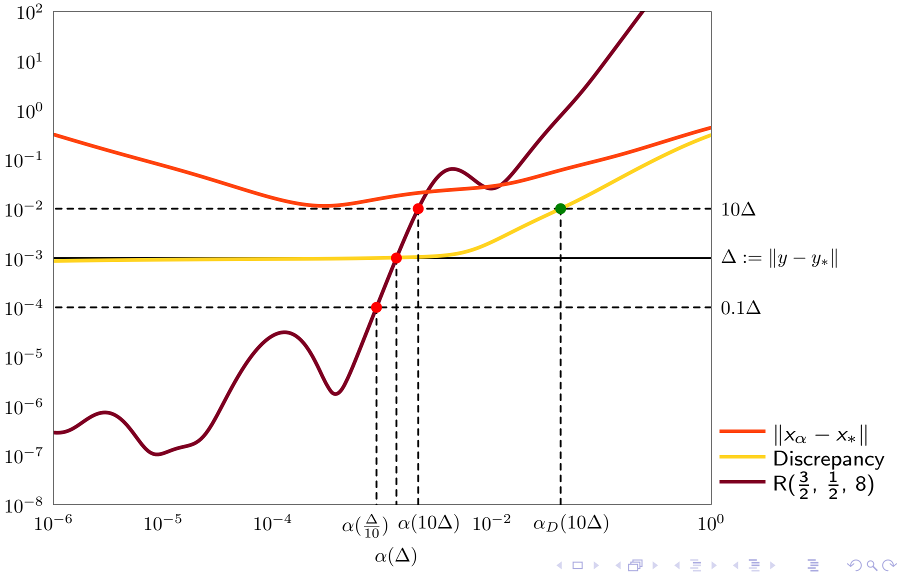
$$\|x_\alpha - x_*\| \leq C(b) \inf_{\alpha \geq 0} \left\{ \|x_\alpha^+ - x_*\| + \frac{\delta}{2\sqrt{\alpha}} \right\},$$

where x_α^+ is the approximate solution with exact right-hand side. It holds $\sup_{\|y-y_*\|\leq\delta} \|x_\alpha - x_\alpha^+\| \leq \frac{\delta}{2\sqrt{\alpha}}$

- Largest solution $\Rightarrow$ stability
- Smallest solution $\Rightarrow$ quasi-optimality
- If the solution is unique, quasi-optimality also holds for the largest solution. In most of our numerical experiments the solution was unique.

In the following we choose the largest solution.

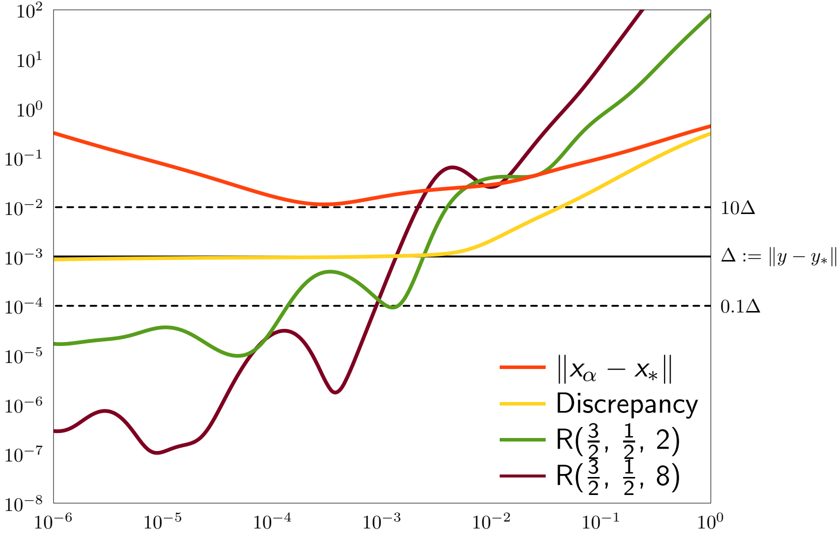
Stability of choice $\alpha = \alpha(\delta)$ from rule $d(\alpha) = \delta$



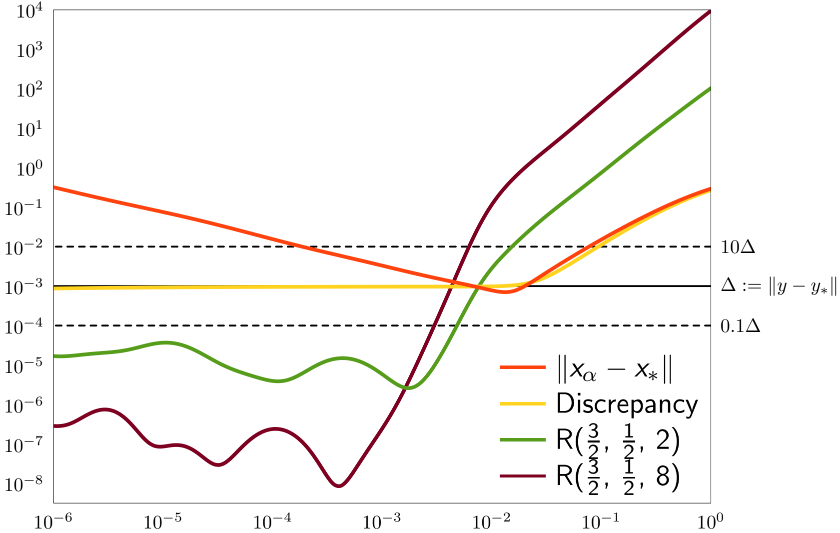
Stability of parameter choice

- Compare rules for choice of the regularization parameter $\alpha = \alpha(\delta)$ as the solution of the equation $d(\alpha) = b\delta$.
- The stability of parameter choice rule with respect to the inaccuracy of noise level information increases for increasing $d'(\alpha)$ in the neighbourhood of $\alpha(\|y - y_*\|)$.
- In many rules from the family $d'(\alpha)$ is much larger than in the discrepancy principle, thus these rules are more stable with respect to inaccuracies of noise level $\delta \approx \|y - y_*\|$.
- The previous slide and the following 3 slides show the behaviour of functions $d(\alpha)$ in the problem 'phillips' from Hansen's Regularization Tools.

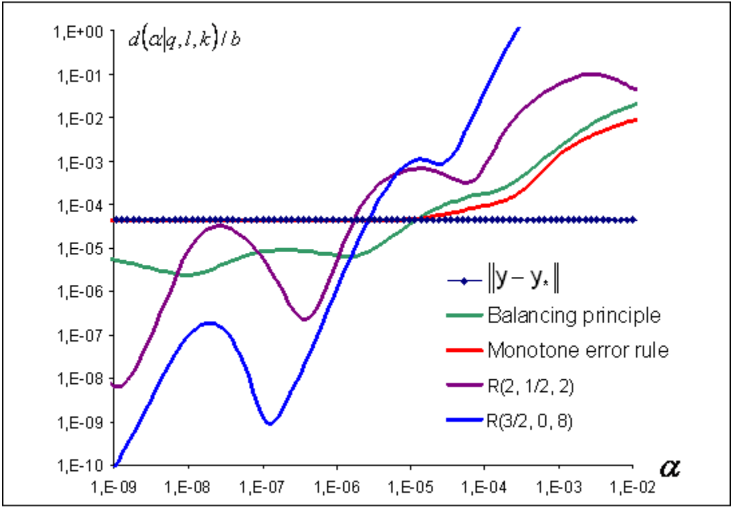
Behavior of functions $d(\alpha)$ in rules $d(\alpha) = \delta$, $p = 0$



Behavior of functions $d(\alpha)$ in rules $d(\alpha) = \delta$, $p = 2$



Behaviour of function $d(\alpha)$ in the neighbourhood $\alpha(\|y - y_*\|)$, $p = 0$



Hansen's test problems used in numerical tests.

Set I of test problems, P. C. Hansen's *Regularization Tools*.

Nr	Problem	cond_{100}	selfadj	Description
1	baart	5e+17	no	(Artificial) Fredholm integral equation of the first kind
2	deriv2	1e+4	yes	Computation of the second derivative
3	foxgood	1e+19	yes	A problem that does not satisfy the discrete Picard condition
4	gravity	3e+19	yes	A gravity surveying problem
5	heat	2e+38	no	Inverse heat equation
6	ilaplace	9e+32	no	Inverse Laplace transform
7	phillips	2e+6	yes	An example problem by Phillips
8	shaw	5e+18	yes	An image reconstruction problem
9	spikes	3e+19	no	Test problem whose solution is a pulse train of spikes
10	wing	1e+20	no	Fredholm integral equation with discontinuous solution

Brezinski-Rodriguez-Seatzu problems

Set II of test problems, *Numerical Algorithms* 2008, 49, 1–4, pp 85–104.

Nr	Problem	cond_{100}	selfadj	Description
11	gauss	6e+18	yes	Test problem with Gauss matrix $a_{ij} = \sqrt{\frac{\pi}{2\sigma}} e^{-\frac{\sigma}{2(i-j)^2}}$, kus $\sigma = 0.01$
12	hilbert	4e+19	yes	Test problem with Hilbert matrix $a_{ij} = \frac{1}{i+j-1}$
13	lotkin	2e+21	no	Test problem with Lotkin matrix (same as Hilbert matrix, except $a_{1j} = 1$)
14	moler	2e+4	yes	Test problem with Moler matrix $A = B^T B$, where $b_{ii} = 1$, $b_{i,i+1} = 1$, and $b_{ij} = 0$ otherwise
15	pascal	1e+60	yes	Test problem with Pascal matrix $a_{ij} = \binom{i+j-2}{i-1}$
16	prolate	1e+17	yes	Test problem with a symmetric, ill-conditioned Toeplitz matrix

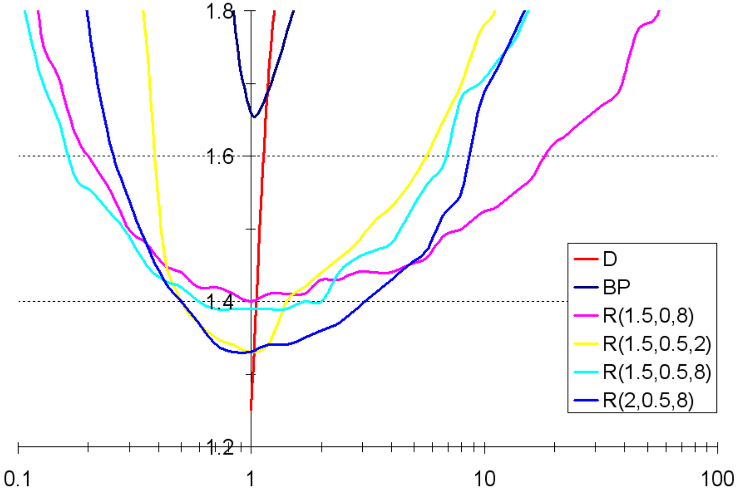
Solution vectors for BRS-problems

Description	$\overline{x}_i$
constant	1
linear	$\frac{i}{N}$
quadratic	$\left(\frac{i-\lfloor\frac{N}{2}\rfloor}{\lceil\frac{N}{2}\rceil}\right)^2$
sinusoidal	$\sin\frac{2\pi(i-1)}{N}$
linear+sinusoidal	$\frac{i}{N}+\frac{1}{4}\sin\frac{2\pi(i-1)}{N}$
step function	$\begin{cases}0, & \text{if } i\leq\lfloor\frac{N}{2}\rfloor\\ 1, & \text{if } i>\lfloor\frac{N}{2}\rfloor\end{cases}$

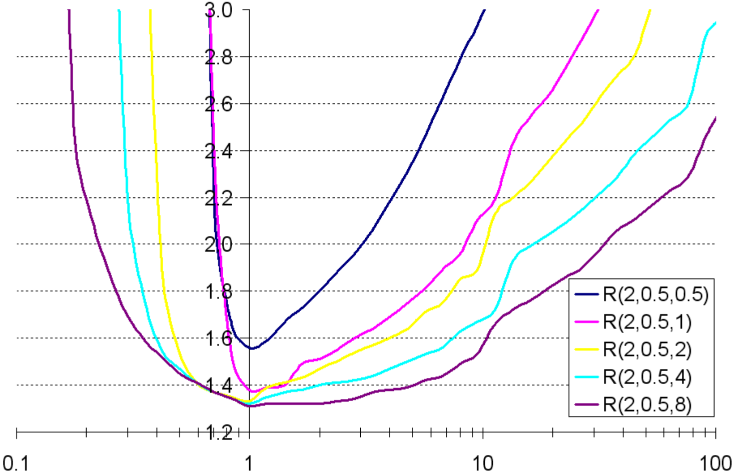
Perturbed data and presentation of results

- Besides solution x_* also smoother solution $x_{*,p} = (A^*A)^{p/2}x_*$ with $y_* = Ax_{*,p}$, $p = 2$ was used.
- The problems were normalized, so that Euclidean norms of the operator and the right hand side were 1.
- For perturbed data we took $y = y_* + \Delta$, $\|\Delta\| = 0.3,\ 10^{-1},\ \ldots,\ 10^{-6}$ with 10 different normally distributed perturbations Δ generated by computer.
- Problems were solved by Tikhonov method, assuming that the noise level is $\delta = \varrho\|y - y_*\|$. Thus $\varrho > 1$ corresponds to overestimation of the true error, $\varrho < 1$ to underestimation.
- To compare the rules, we present averages (over problems, perturbations Δ and runs) of error ratios $\|x_\alpha - x_*\|/e_{\text{opt}}$, where e_{opt} is minimal error in Tikhonov method.

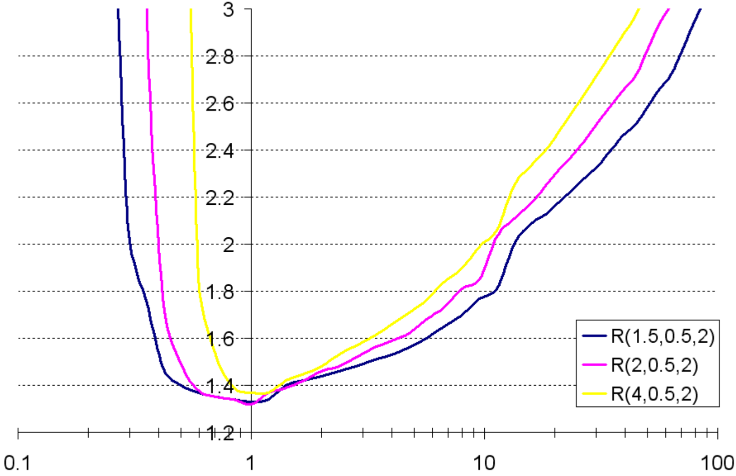
Stability of the rules with respect to $\varrho = \frac{\delta}{\|y-y_*\|}$



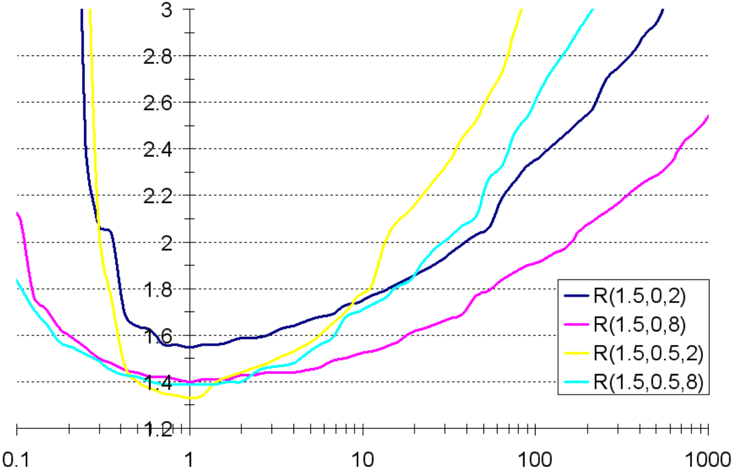
Stability of rule R(q, l, k) increases if k increases



Stability of rule R(q, l, k) increases if q decreases



$l = 0.5$ is recommended ($l = 0$ is good if $\delta \gg \|y - y_*\|$)



Post-estimation of regularization parameter in case $\|y - y_*\| \le \delta$

- $\alpha_{\mathrm{ME}} \ge \alpha_{\mathrm{opt}} := \operatorname{argmin}\{\|x_\alpha - x_*\|,\ \alpha \ge 0\}$, computations suggest $\alpha_{\mathrm{MEe}} = 0.4\alpha_{\mathrm{ME}}$, if $\|y - y_*\| = \delta$.
- More stable with respect to overestimation of noise level is the choice $\alpha_{\mathrm{Me}} = \min(\alpha_{\mathrm{MEe}}, 1.4\alpha_{\mathrm{R}(\frac{3}{2},\ \frac{1}{2},\ 2)})$, $b = 0.023$.

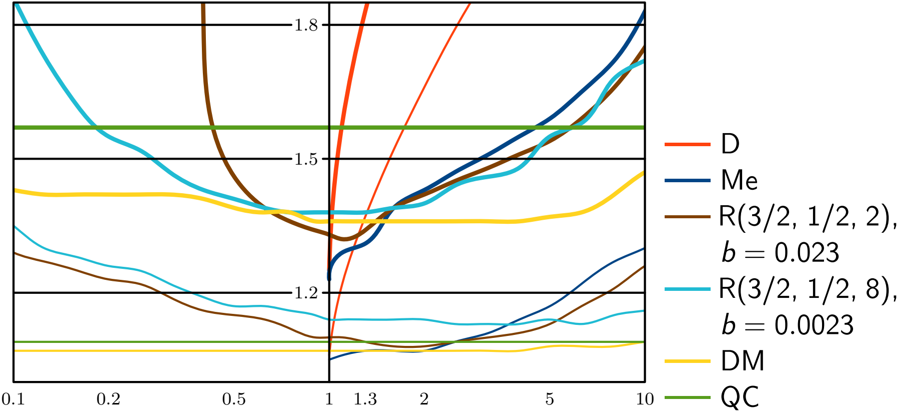
Heuristic rules (not using δ) in Tikhonov method

- **Quasioptimality criterion Q**: take α as the global minimizer of the function $\psi(\alpha) = \|x_\alpha - x_{2,\alpha}\|$, where $x_{2,\alpha}$ is 2-iterated Tikhonov approximation $x_{2,\alpha} = (\alpha I + A^*A)^{-1}(\alpha x_\alpha + A^*y)$. Sometimes this gives too small α, therefore we try to find a lower bound of minimization interval, determined during computations.
- **Rule QC**. Make computations on the sequence of parameters $\alpha_i = q^{i-1}$, $i = 1, 2, \ldots;$ $q < 1$, for example, $q = 0.9$. Take α_i as the minimizer of the function $\psi(\alpha_i) = \|x_{\alpha_i} - x_{2,\alpha_i}\|$ in the interval $[\underline{\alpha}, 1]$, where $\underline{\alpha}$ is the largest α_i, for which the value of $\psi(\alpha_i)$ is $C = 5$ times larger than its value at its current minimum.
- L-curve rule, GCV-rule, Hanke-Raus rule and Brezinski-Rodriguez-Seatzu rule gave in our numerical experiments not so good results as rules Q and QC.

Rule DM for approximate noise level in Tikhonov method

- **Rule DM for Tikhonov method**
  1) Make computations on the sequence of parameters $\alpha_i = q^{i-1}$, $i = 1, 2, \ldots$; $q < 1$, for example, $q = 0.9$; find $\underline{\alpha}$ as the first α_i for which $\sqrt{\alpha_i}\|x_{\alpha_i} - x_{2,\alpha_i}\| \leq c_1\delta$, $c_1 = \text{const}$;
  2) find $\alpha_i = \operatorname{argmin} \frac{(1+\alpha\|A\|^{-2})\|D_\alpha^{1/2}B_\alpha(Ax_\alpha - y)\|^2}{\alpha^{c_2}\|D_\alpha^{1/2}B_\alpha^2(Ax_\alpha - y)\|}$ in $[\underline{\alpha}, 1]$, $c_2 = \text{const}$.
- If $\varrho := \delta/\|y - y_*\| \in (0.1, 10)$, then we recommend $c_1 = 0.005$, $c_2 = 0.05$; if less information is known, $\varrho \in (0.01, 100)$, then we recommend $c_1 = 0.001$, $c_2 = 0.47$.
- Convergence $x_\alpha \to x_*$, as $\delta \to 0$, provided that $\lim \|y - y_*\|/\delta \leq C$, is guaranteed. If $x_* \in \mathcal{R}((A^*A)^{p/2})$, then for rule DM with $c_1 \geq 0.24$ the error estimate $\|x_\alpha - x_*\| \leq \text{const}\, \delta^{p/(p+1)}$ holds for all $p \leq 2$.

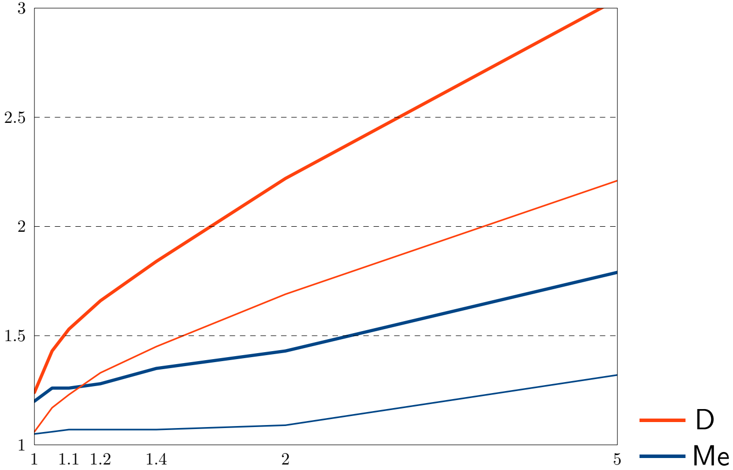
Averages (thick lines) and medians (thin lines) of error ratios in various rules in dependence of $\varrho = \delta / \|y - y_*\|$



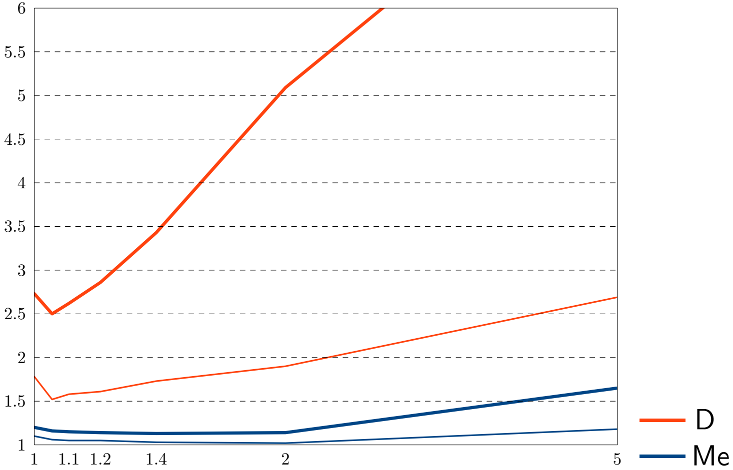
Preferences of rules in dependence of the accuracy of noise level information $\varrho = \delta / \|y - y_*\|$

- If we are sure that $\varrho \in [1, 1.5]$, then we recommend the rule Me.
- In case $\varrho \in [0.6, 1.5]$ we recommend the rule R(3/2, 1/2, 2), $b = 0.023$.
- If less information about the noise level is known, for example, $\varrho \in [1/20, 20]$, then we recommend the rule DM.
- For even less information about the noise level, we recommend the rule QC. If $\|Ax_{\alpha_{\mathrm{QC}}} - y\|$ is evidently less than $\|y - y_*\|$, then we recommend to decrease the constant C, for example, using $(C+1)/2$ instead of C.

Averages (thick lines) and medians (thin lines) of error ratios in rules D and Me in dependence of $\varrho = \delta / \|y - y_*\|$, $p = 0$



Averages (thick lines) and medians (thin lines) of error ratios in rules D and Me in dependence of $\varrho = \delta/\|y - y_*\|$, $p = 2$



Conclusions

- We propose a family of rules R(q, l, k) for approximate noise level, where $3/2 \le q < \infty$, $l \ge 0$, $k \ge l/q$, $2q, 2k, 2l \in \mathbb{N}$.
- If $k > l/q$ and $\frac{\|y - y_*\|}{\delta} \le C = \text{const}$ as $\delta \to 0$, then we have $\|x_\alpha - x_*\| \to 0$ $(\delta \to 0)$.
- Certain rules from the family gave in numerical experiments good results in case of several times over- or underestimated noise level.

Bibliography

1. U. Hämarik, R. Palm, and T. Raus. On minimization strategies for choice of the regularization parameter in ill-posed problems. *Numerical Functional Analysis and Optimization*, 30(9&10):924–950, 2009.
2. U. Hämarik and T. Raus. About the balancing principle for choice of the regularization parameter. *Numerical Functional Analysis and Optimization*, 30(9&10):951–970, 2009.
3. T. Raus and U. Hämarik. New rule for choice of the regularization parameter in (iterated) Tikhonov method. *Mathematical Modelling and Analysis*, 14(2):187–198, 2009.
4. R. Palm. *Numerical comparison of regularization algorithms for solving ill-posed problems*. PhD thesis, University of Tartu, 2010. `http://hdl.handle.net/10062/14623`.
5. U. Hämarik, R. Palm, and T. Raus. Comparison of parameter choices in regularization algorithms in case of different information about noise level. *Calcolo*, 48(1):47–59, 2011.
6. U. Hämarik, R. Palm, T. Raus. A family of rules for parameter choice in Tikhonov regularization of ill-posed problems with inexact noise level. *Journal of Computational and Applied Mathematics*, 2011. Accepted. doi:10.1016/j.cam.2011.09.037